



Serving Those Who Serve Our Country



Data Science at the Defense Personnel and Security Research Center

Mission: Improve the Effectiveness, Efficiency, and Fairness of DoD Personnel Security and Suitability Programs

**Eric L. Lang, Ph.D., Director, PERSEREC
Defense Personnel and Security Research Center (PERSEREC)
Defense Manpower Data Center (DMDC)**

May 14, 2015

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the Department of Defense, the Defense Manpower Data Center, or the Defense Personnel and Security Research Center

Defense Personnel and Security Research Center



PERSEREC

UNCLASSIFIED

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 14 MAY 2015		2. REPORT TYPE		3. DATES COVERED 00-00-2015 to 00-00-2015	
4. TITLE AND SUBTITLE Data Science at the Defense Personnel and Security Research Center				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Defense Manpower Data Center (DMDC), Defense Personnel and Security Research Center (PERSEREC), Alexandria, VA				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES Presented at the 12th Annual Acquisition Research Symposium held May 13-14, 2015 in Monterey, CA.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 9	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

PERSEREC – Who We Are

1986

In the wake of catastrophic espionage by John Walker and other spies in the mid 1980s, PERSEREC was established to improve the effectiveness, efficiency, and fairness of personnel security in the DoD

Since 2011

In the wake of events like 9-11, Ft. Hood, and the Washington Navy Yard shootings, we expanded our mission to address a broader range of human capital, insider threat, workplace safety, and security issues

Throughout our history, we have:

- Conducted long-term programmatic research for the military, security, and intelligence communities
- Provided quick-response analyses and studies in support of policy makers, systems operations, military leaders, and security practitioners
- Developed innovative tools and job aids

PERSEREC's Core Capabilities

1. Automated Background Record Checks
2. Emerging Indicators, e.g., Social Media, Mental Health
3. Special Policy Initiatives, e.g., Suicide & Sexual Assault
4. Security and Suitability Vetting
5. Insider Threat
6. **Data Science**

Several current and planned initiatives apply advanced data integration and analytic techniques to improve the quality, speed and cost-effectiveness of solutions to important policy and operational challenges.

Data Science: Current Initiatives

- Data Science Hadoop cluster projects since 2013, e.g.:
 - ✓ Modeling attrition and retention of high value mil occupations
 - ✓ Exploring innovative efficiencies for DMDC database operations
 - ✓ Modeling security clearance requirements for OUSD(I)
- Working on wider DoD Data Science Collaborations, e.g.:
 - ✓ DMDC as the Advanced Analytics Center for OUSD(P&R)
 - ✓ Partnerships with NPS, local universities and other groups

Data Science: Possible AT&L Initiatives

1. Modeling/Predicting Key Acquisition Program Outcomes
Goal: Identify key and unknown causal factors associated with the performance of AT&L programs and contracts.
 - ✓ Predictors: structured and natural language data from AT&L DBs
 - ✓ Analysis: e.g., Random Forest and Logistic Regression
 - ✓ Predicting: Factors associated with contracts/programs being:
 - on-time (or length of over-run)
 - within budget (or amount over/under budget)
 - acceptable deliverables (ideally quality of products/services)

Data Science: Possible AT&L Initiatives

2. Modeling/Estimating DoD Security Costs

Goal: Create a taxonomy of DoD security types and costs to help estimate overall costs of security across DoD.

- ✓ Predictors: e.g., DTIC, FPDS, DAMIR, personnel records, others
- ✓ Analysis: e.g., Clustering and taxonomic modeling
- ✓ Predicting: Taxonomies and costs for each taxonomy, covering:
 - Personnel security
 - Information security
 - Physical security
 - Industrial security
 - Operations security

Summary Thoughts

- Potential efficiencies in AT&L's \$360B managed portfolio
- Extant Data + Data Science = innovative actionable insights
- Data Science HW & SW are necessary but not sufficient
- Success = Data Science HW & SW + domain expertise
- Data Science capabilities exist now in the Monterey area

Back Up Slides

PERSEREC: Example Areas of Success

Success = Implementation of PERSEREC recommendations and tools

- Production and revision of the National Adjudicative Guidelines specified in E.O. 12968
- Phased Periodic Reinvestigation (PPR) method for continuing TOP SECRET access eligibility
- Electronic adjudication (“eAdjudication”) of clean SECRET cases
- Tools for evaluating the quality of investigations and adjudications
 - RAISE: Rapid Assessment of Incomplete Security Evaluations
 - RADAR: Review of Adjudication Documentation Accuracy and Rationales
- Automated Continuous Evaluation System (ACES) for clearances holders IOC
- Tools for security and intel professionals (e.g., “Adjudication Desk Reference,” “Customizable Security Guide,” and “Insider Risk Evaluation and Audit Tool”)
- Espionage trend analyses (e.g., “Changes in Espionage by Americans 1947-2007”)
- DIRE: “Dispositional Indicators of Risk Exposure” [DIRE] tool for assessing personality characteristics that are security and safety risks
- DoD adjudicator certification standards and procedures

The PPR and eAdjudication have saved the DoD more than **\$250M**